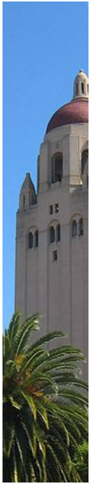


Artificial Intelligence and Economic Growth

P. Aghion, B. Jones, and C. Jones

in Agrawal et al *The Economics of Artificial Intelligence*, 2019

What are the implications of A.I. for economic growth?

- Build some growth models with A.I.
  - A.I. helps to make goods
  - A.I. helps to make ideas
- Implications
  - Long-run growth
  - Share of GDP paid to labor vs capital
  - Firms and organizations
- Singularity?

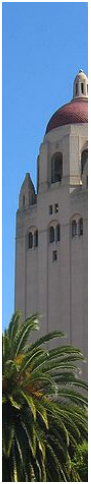
Two Main Themes

- A.I. modeled as a continuation of automation
  - Automation = replace labor in particular tasks with machines and algorithms
  - *Past:* textile looms, steam engines, electric power, computers
  - *Future:* driverless cars, paralegals, pathologists, maybe researchers, maybe everyone?

- A.I. may be limited by Baumol's cost disease
  - *Baumol:* growth constrained not by what we do well but rather by what is essential and yet hard to improve

Outline

- Basic model: automating tasks in production
- A.I. and the production of new ideas
- Singularity?
- Some facts



The Zeira 1998 Model

Simple Model of Automation (Zeira 1998)

- Production uses n tasks/goods:

$$Y = AX_1^{\alpha_1} X_2^{\alpha_2} \cdot \ldots \cdot X_n^{\alpha_n},$$

where $\sum_{i=1}^{n} \alpha_i = 1$ and

$$X_{it} = \begin{cases} L_{it} & \text{if not automated} \\ K_{it} & \text{if automated} \end{cases}$$

- Substituting gives

$$Y_t = A_t K_t^{\alpha} L_t^{1-\alpha}$$

$$Y_t = A_t K_t^{\alpha} L_t^{1-\alpha}$$

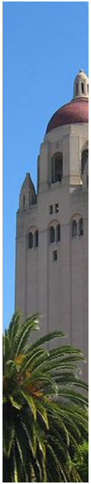
- Comments:
  - α reflects the *fraction* of tasks that are automated
  - Embed in neoclassical growth model $\Rightarrow$

$$g_y = \frac{g_A}{1-\alpha} \quad \text{where} \quad y_t \equiv Y_t / L_t$$

- Automation: $\uparrow \alpha$ raises both capital share and LR growth
  - Hard to reconcile with 20th century
  - Substantial automation but stable growth and capital shares

Subsequent Work

- Acemoglu and Restrepo (2017, 2018, 2019, 2020, 2021, ...)
  - Old tasks are gradually automated as new (labor) tasks are created
  - Fraction automated can then be steady
  - Rich framework, with endogenous innovation and automation, all cases worked out in great detail
- Peretto and Seater (2013), Hemous and Olson (2016), Agrawal, McHale, and Oettl (2017)



Automation and Baumol’s Cost Disease

Baumol's Cost Disease and the Kaldor Facts

- Baumol: Agriculture and manufacturing have rapid growth and declining shares of GDP
  - ... but also rising automation
- Aggregate capital share could reflect a balance
  - Rises within agriculture and manufacturing
  - But falls as these sectors decline
- Maybe this is a general feature of the economy!
  - First agriculture, then manufacturing, then services

AJJ Economic Environment

Final good $$Y_t = \left(\int_0^1 X_{it}^{\frac{\sigma-1}{\sigma}} \, di \right)^{\frac{\sigma}{\sigma-1}} \quad \text{where} \quad \sigma < 1$$

Tasks $$X_{it} = \begin{cases} K_{it} & \text{if automated} \quad i \in [0, \beta_t] \\ L_{it} & \text{if not automated} \quad i \in [\beta_t, 1] \end{cases}$$

Capital accumulation $$\dot{K}_t = I_t - \delta K_t$$

Resource constraint (K) $$\int_0^1 K_{it} di = K_t$$

Resource constraint (L) $$\int_0^1 L_{it} di = L$$

Resource constraint (Y) $$Y_t = Cons_t + I_t$$

Allocation $$I_t = \bar{s}_K Y_t$$

Automation and growth

- Combining equations

$$Y_t = \left[\beta_t \left(\frac{K_t}{\beta_t} \right)^{\frac{\sigma-1}{\sigma}} + (1-\beta_t) \left(\frac{L}{1-\beta_t} \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}$$

- How β interacts with K: two effects
  - β: what fraction of tasks have been automated
  - β: Dilution as $K/\beta \Rightarrow K$ spread over more tasks
- Same for labor: $L/(1-\beta_t)$ means given L concentrated on fewer tasks, raising "effective labor"

Rewriting in classic CES form

- Collecting the β terms into factor-augmenting form:

$$Y_t = F(B_t K_t, C_t L_t)$$

where

$$B_t = \left(\frac{1}{\beta_t}\right)^{\frac{1}{1-\sigma}} \quad \text{and} \quad C_t = \left(\frac{1}{1-\beta_t}\right)^{\frac{1}{1-\sigma}}$$

- Effect of automation: $\uparrow \beta_t \Rightarrow \downarrow B_t$ and $\uparrow C_t$

Intuition: dilution effects just get magnified since $\sigma < 1$

Automation

- Suppose a constant fraction of non-automated tasks get automated every period:

$$\dot{\beta}_t = \theta(1 - \beta_t)$$

$$\Rightarrow \quad \beta_t \to 1$$

- What happens to $1 - \beta_t =: m_t$?

$$\frac{\dot{m}_t}{m_t} = -\theta$$

The fraction of labor-tasks falls at a constant exponential rate

Putting it all together

$$Y_t = F(B_t K_t, C_t L_t) \text{ where } B_t = \left(\frac{1}{\beta_t}\right)^{\frac{1}{1-\sigma}} \text{ and } C_t = \left(\frac{1}{1-\beta_t}\right)^{\frac{1}{1-\sigma}}$$

- $\beta_t \to 1 \Rightarrow B_t \to 1$
- But C_t grows at a constant exponential rate!

$$\frac{\dot{C}_t}{C_t} = -\frac{1}{1-\sigma}\frac{\dot{m}_t}{m_t} = \frac{\theta}{1-\sigma}$$

- When a constant fraction of remaining goods get automated and $\sigma < 1$, the automation model features an asymptotic BGP that satisfies Uzawa

Factor Shares of Income

- Ratio of capital share to labor share:

$$\frac{\alpha_{K_t}}{\alpha_{L_t}} = \left(\frac{\beta_t}{1-\beta_t}\right)^{1/\sigma} \left(\frac{K_t}{L_t}\right)^{\frac{\sigma-1}{\sigma}}$$

- Two offsetting effects ($\sigma < 1$):
  - $\uparrow \beta_t$ raises the capital share
  - $\uparrow K_t/L_t$ lowers the capital share

These balance and deliver constant factor shares in the limit

$$\alpha_{Kt} \equiv \frac{F_K K}{Y} = \beta_t^{\frac{1}{\sigma}} \left(\frac{K_t}{Y_t}\right)^{\frac{\sigma-1}{\sigma}} \to \left(\frac{\bar{s}_K}{g_Y+\delta}\right)^{\frac{\sigma-1}{\sigma}} < 1$$

Intuition for AJJ result

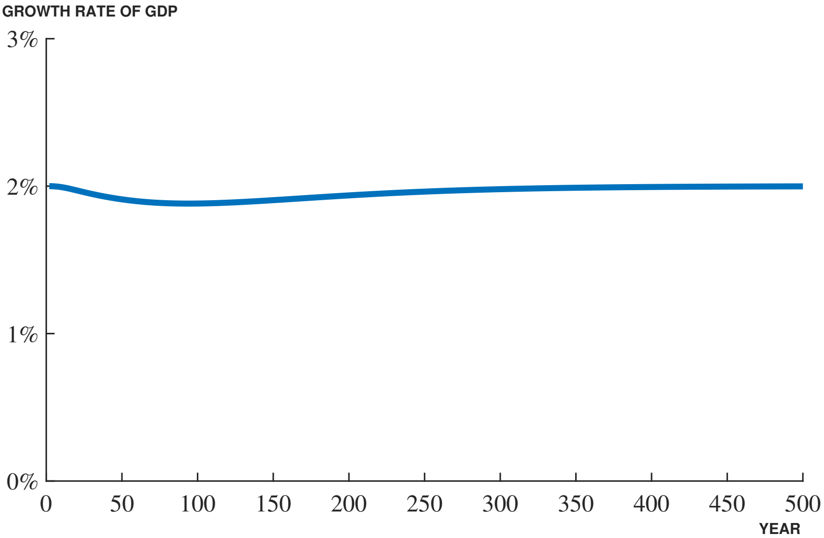
- Why does automation lead to balanced growth and satisfy Uzawa?
  - $\beta_t \to 1$ so the KATC piece "ends" eventually (all tasks automated)
  - Labor per task: $L/(1-\beta_t)$ rises exponentially over time!
  - Constant population, but concentrated on an exponentially shrinking set of goods $\Rightarrow$ exponential growth in "effective" labor

- Baumol logic
  - Agr/Mfg shrink as a share of the economy...
  - Labor still gets 2/3 of GDP! Vanishing share of tasks, but all else is cheap (Baumol)

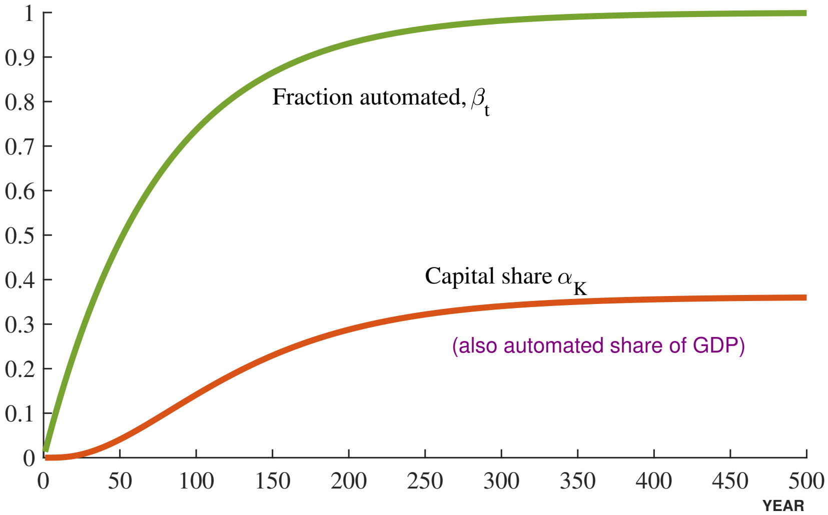
Interesting question: What fraction of tasks automated today? β_{2022}

(B. Jones and X. Liu 2022 on capital-embodied technical change)

Simulation: Automation and Asymptotic Balanced Growth



Simulation: Capital Share and Automation Fraction



Constant Factor Shares?

- Consider $g_A > 0$ — technical change beyond just automation
- Alternatively, factor shares can be constant if automation follows

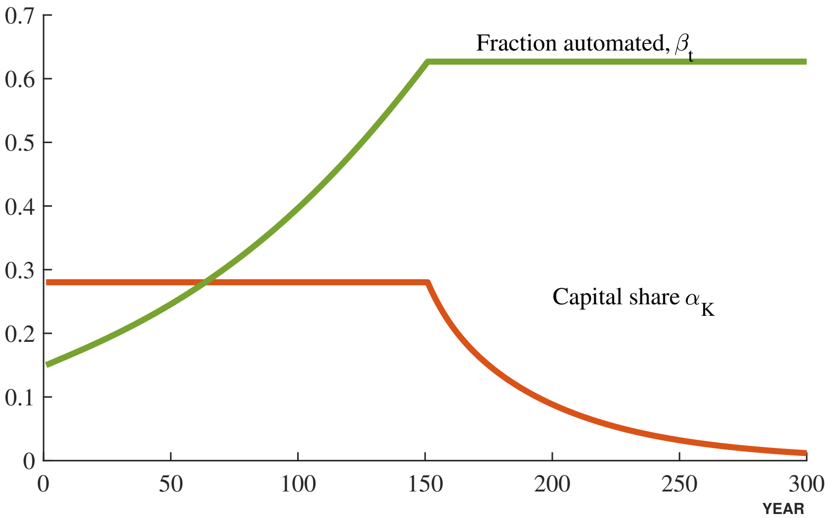
$$g_{\beta t} = (1 - \beta_t) \left(\frac{-\rho}{1 - \rho} \right) g_{kt},$$

- Knife-edge condition...
- Surprise: growth rates increase not decrease. Why? Requires

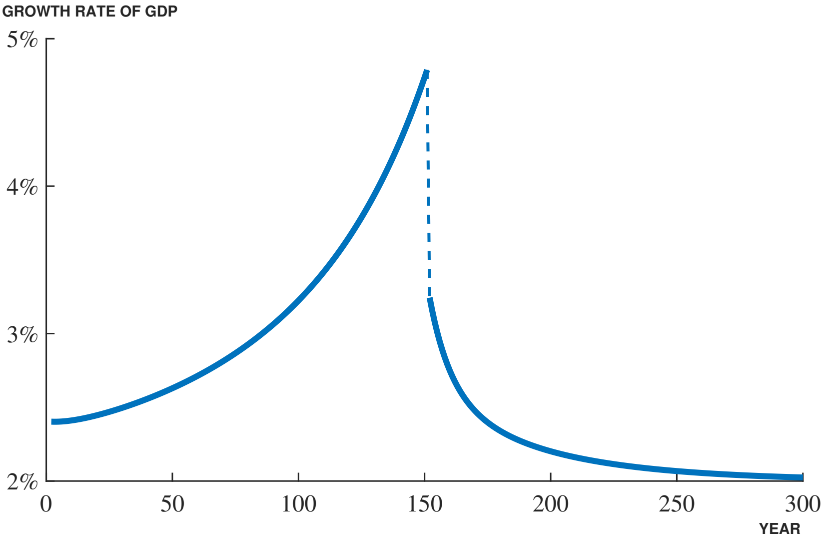
$$g_{Yt} = g_A + \beta_t g_{Kt}.$$

- $g_A = 0$ means zero growth. $g_A > 0$ means growth rises

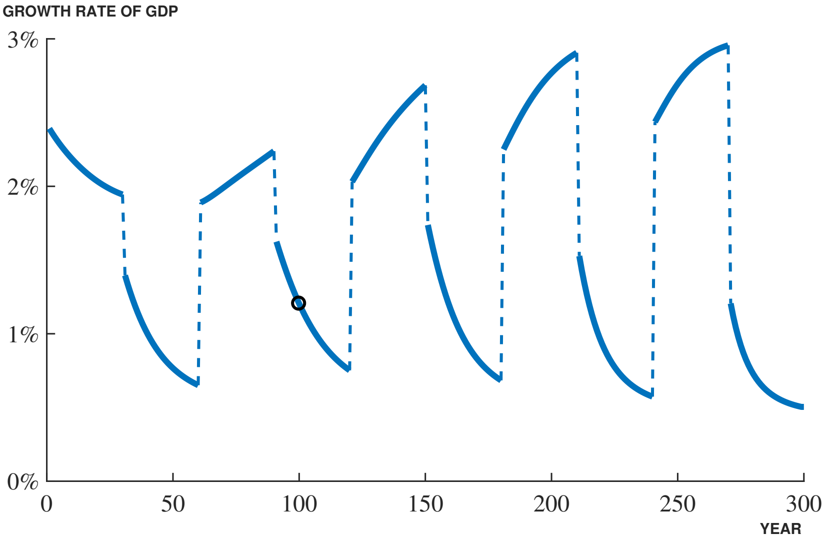
Simulation: Constant Capital Share



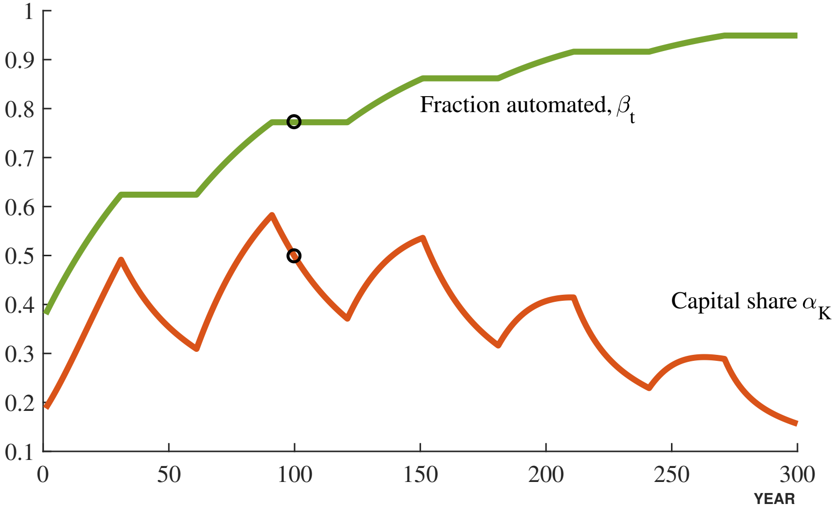
Simulation: Constant Capital Share

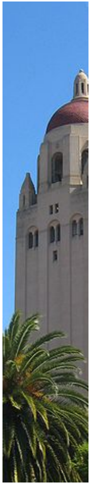


Simulation: Switching regimes...



Simulation: Switching regimes...





A.I. and Ideas

AI in the Ideas Production Function

- Let production of goods and services be $Y_t = A_t L_t$
- Let idea production be:

$$\dot{A}_t = A_t^{\phi} \left(\int_0^1 X_{it}^{\frac{\sigma-1}{\sigma}} di \right)^{\frac{\sigma}{\sigma-1}}, \ \sigma < 1$$

- Assume fraction β_t of tasks are automated by date t. Then:

$$\dot{A}_t = A_t^{\phi} F(B_t K_t, C_t S_t)$$

where

$$B_t = \left(\frac{1}{\beta_t} \right)^{\frac{1}{1-\sigma}} \text{ and } C_t = \left(\frac{1}{1-\beta_t} \right)^{\frac{1}{1-\sigma}}$$

- This is like before...

AI in the Ideas Production Function

- Intuition: with $\sigma < 1$ the scarce factor comes to dominate

$$F(B_t K_t, C_t S_t) = C_t S_t \, F\left(\frac{B_t K_t}{C_t S_t}, 1\right) \to C_t S_t$$

- So, with continuous automation

$$\dot{A}_t \to A_t^{\phi} C_t S_t$$

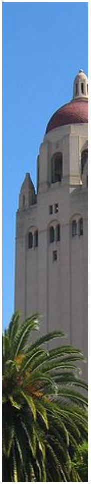
- And asymptotic balanced growth path becomes

$$g_A = \frac{g_C + g_S}{1 - \phi}$$

- We get a "boost" from continued automation (g_C)

Can automation replace population growth?

- Maybe! Suppose S is constant, $g_S = 0$
  - Intuition: Fixed S is spread among exponentially-declining measure of tasks
  - So researchers per task is growing exponentially!
- However
  - This setup takes automation as exogenous and at "just the right rate"
  - What if automation is endogenized?
  - Is population growth required to drive automation?
  - Could a smart/growing AI entirely replace humans?



Singularities

Singularities

- Now we become more radical and consider what happens when we go "all the way" and allow AI to take over all tasks.
- **Example 1:** Complete automation of goods and services production.

$$Y_t = A_t K_t$$

$\rightarrow$ Then growth rate can accelerate exponentially

$$g_Y = g_A + sA_t - \delta$$

we call this a "Type I" growth explosion

Singularities: Example 2

- Complete automation in ideas production function

$$\dot{A}_t = K_t A_t^{\phi}$$

- Intuitively, this idea production function acts like

$$\dot{A}_t = A_t^{1+\phi}$$

- Solution:

$$A_t = \left(\frac{1}{A_0^{-\phi} - \phi t} \right)^{1/\phi}$$

- Thus we can have a true **singularity** for $\phi > 0$. A_t exceeds any finite value before date $t^* = \frac{1}{\phi A_0^{\phi}}$.

Singularities: Example 3 – Incomplete Automation

- Cobb-Douglas, α and β are fraction automated, S constant

$$\dot{K}_t = \bar{s}L^{1-\alpha}A_t^{\sigma}K_t^{\alpha} - \delta K_t.$$

$$\dot{A}_t = K_t^{\beta}S^{\lambda}A_t^{\phi}$$

- Standard endogenous growth requires $\gamma = 1$:

$$\gamma := \frac{\sigma}{1-\alpha} \cdot \frac{\beta}{1-\phi}.$$

- If $\gamma > 1$, then growth explodes!
  - Can occur without full automation
  - Example: $\alpha = \beta = \phi = 1/2$ and $\sigma > 1/2$.

Objections to singularities

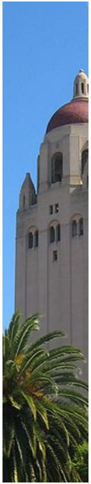
1. Automation limits (no $\beta_t \to 1$)
2. Search limits

$$\dot{A}_t = A_t^{1+\phi} \quad \text{or even} \quad A_t \leq \bar{A}$$

   but $\phi < 0$ (e.g., fishing out, burden of knowledge...)
3. Natural Laws

$$Y_t = \left(\int_0^1 (a_{it} Y_{it})^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \quad \text{where } \sigma < 1$$

   now can have $a_{it} \to \infty$ for many tasks but no singularity
   - *Baumol theme:* growth determined not by what we are good at, but by what is essential yet hard to improve



Final Thoughts

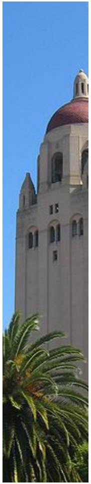
Conclusion: A.I. in the Production of Goods and Services

- Introduced Baumol's "cost disease" insight into Zeira's model of automation
  - Automation can act like labor augmenting technology (surprise!)
  - Can get balanced growth with a constant capital share well below 100%, even with nearly full automation

Conclusion: A.I. in the Ideas Production Function

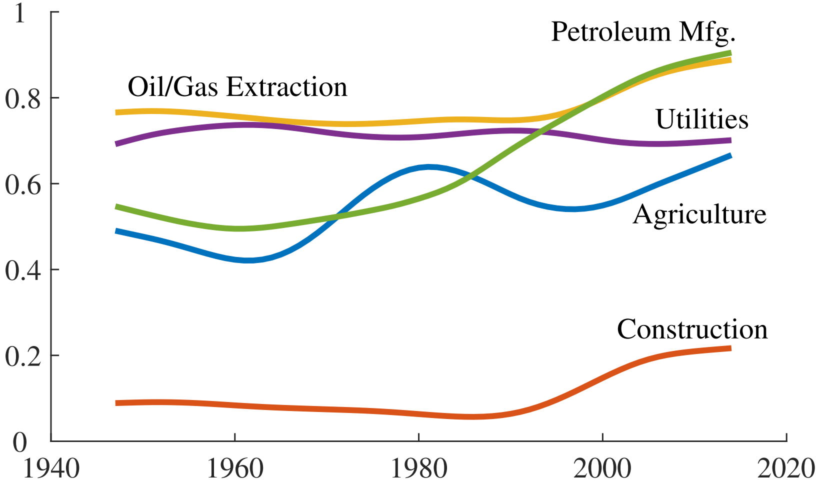
- Could A.I. obviate the role of population growth in generating exponential growth?
- Discussed possibility that A.I. could generate a singularity
  - Derived conditions under which the economy can achieve infinite income in finite time
- Discussed obstacles to such events
  - Automation limits, search limits, and/or natural laws (among others)

Extra Slides

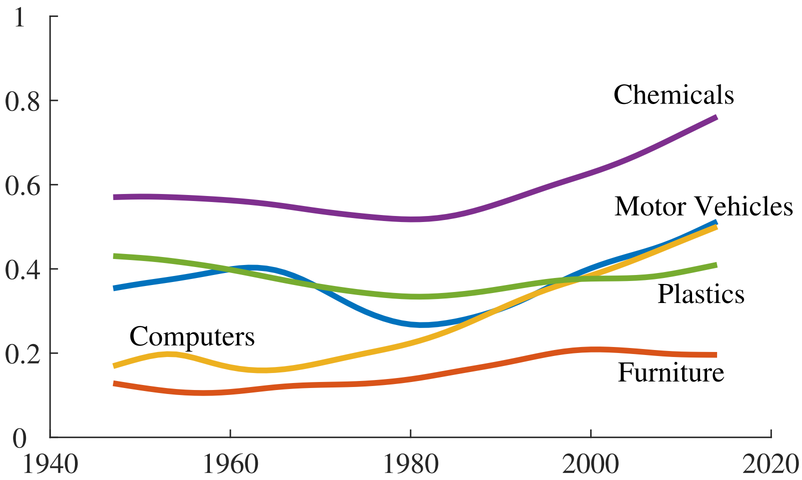


Some Facts

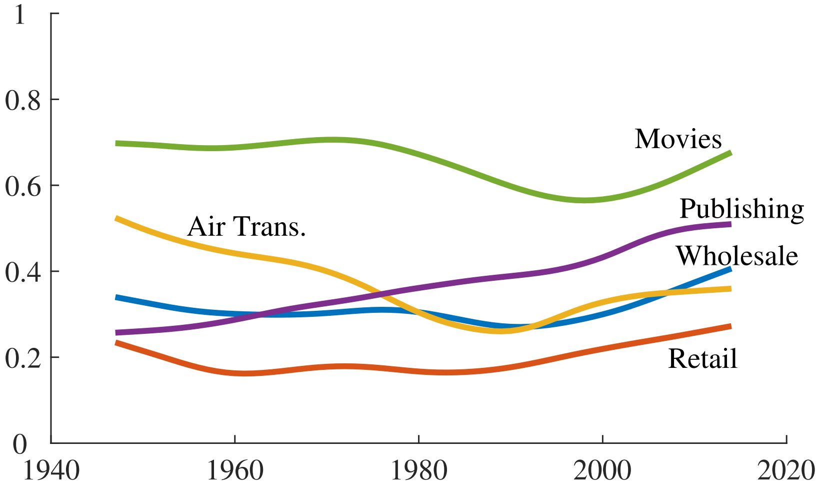
Capital Shares in U.S. Industries



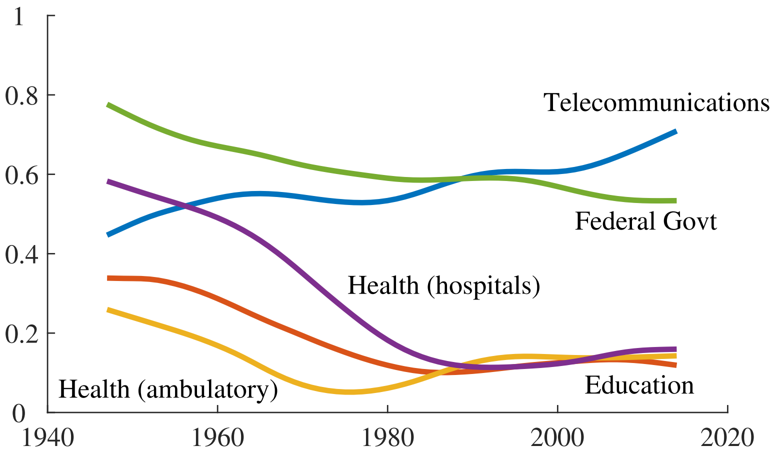
Capital Shares in U.S. Industries



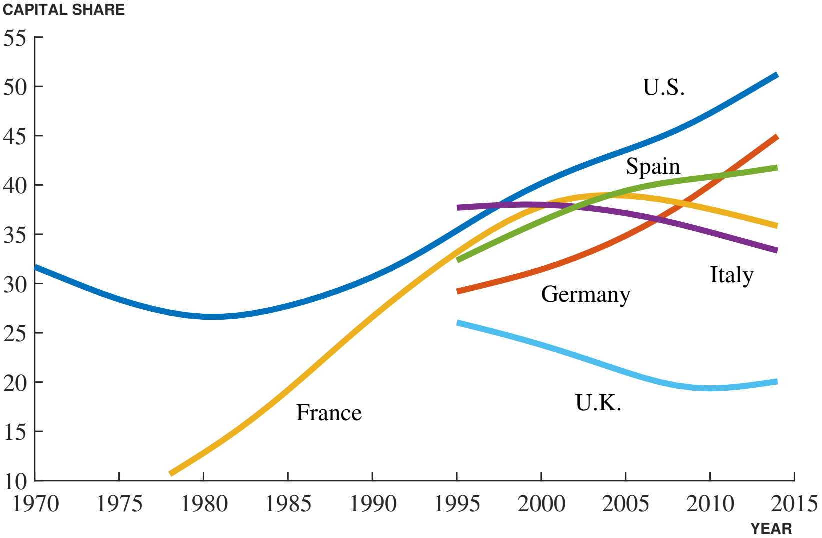
Capital Shares in U.S. Industries



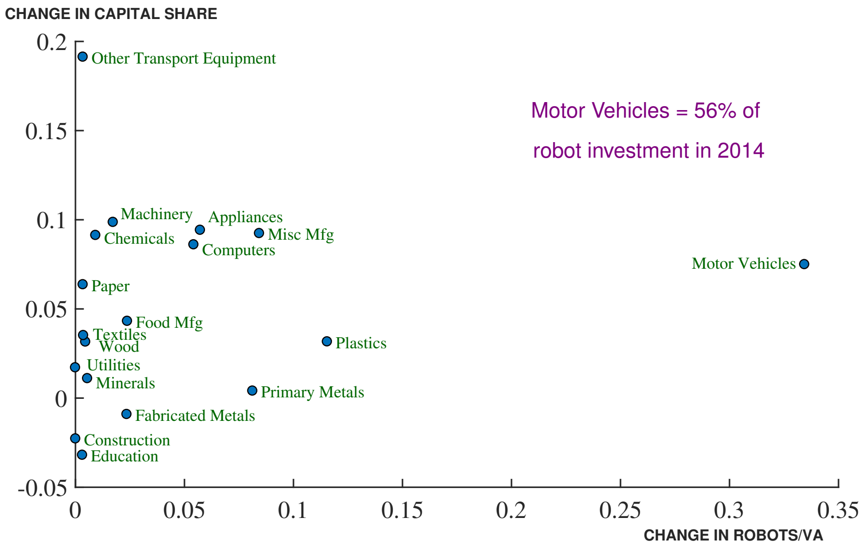
Capital Shares in U.S. Industries



Capital Share of Income: Transportation Equipment



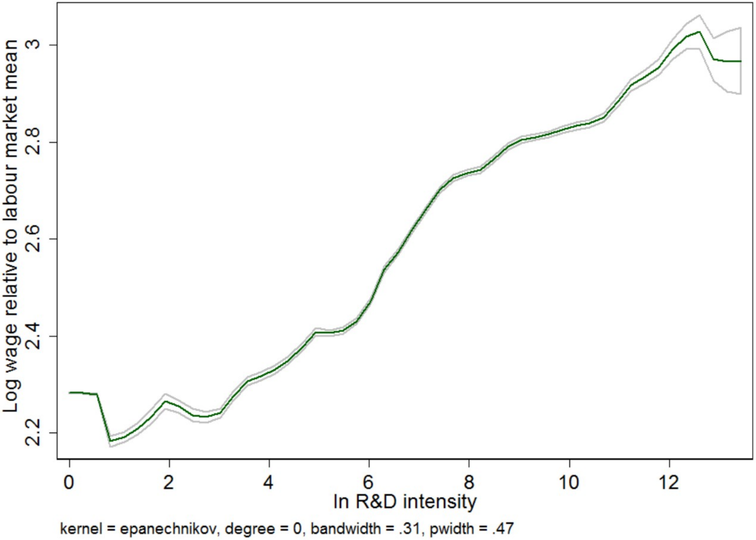
Adoption of Robots and Change in Capital Share



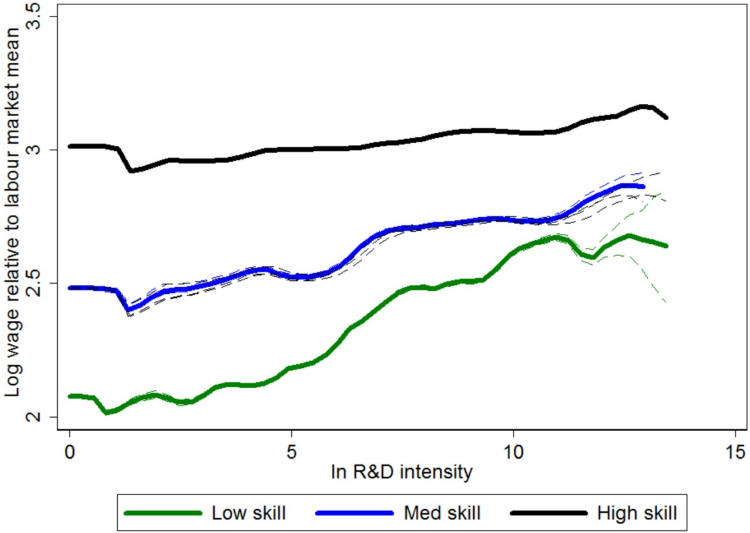
AI, Organizations, and Wage Inequality

- Usual story: robots replace low-skill labor, hence $\uparrow$ skill premium (e.g., Krusell et al. 2000)
- But solving future problems, incl. advancing AI, might be increasingly hard, suggesting $\uparrow$ complementarities across workers, $\uparrow$ teamwork, and changing firm boundaries (Garicano 2000, Jones 2009)
- Aghion et al. (2017) find evidence along these lines
  - outsouce higher fraction of low-skill workers
  - pay *increased* premium to low-skill workers kept

AI, Organizations, and Wage Inequality



AI, Skills, and Wage Inequality



Nonrivalry and the Economics of Data

Chad Jones and Christopher Tonetti

INET-IMF

Macroeconomics in the Age of AI

17 March 2020

Examples of Data

- Google, Facebook
- Amazon
- Tesla, Uber, Waymo
- Medical and genetic data
- Location history
- Speech records
- Physical action data

What is Data in this Paper?

- Data as a factor of production
- Data improves the quality of a product
  - We do not model data as helping a consumer or firm make a more informed decision (e.g., consumption, pricing)
- Data can be useful even if anonymous
- Other aspects of the economics of data are interesting (price discrimination, product specialization, etc.), but are purposely left out of the model

Canonical example: data as input into machine learning algorithm. E.g., medical detection algorithms, self-driving cars, voice recognition software.

Policies on Data Are Being Written Now

What policies governing data use maximize welfare?

- European General Data Protection Regulation (GDPR)
  - Privacy vs. social gain from sharing
  - “The protection of natural persons in relation to the processing of personal data is a fundamental right”
  - “The right. . . must be considered in relation to its function in society. . . ”
- The California Consumer Privacy Act of 2018 (start Jan 1 2020)
  - Allows consumers to opt out of having their data sold
- US Congress: COPRA, ACESS, etc.
- India’s Personal Data Protection bill

Data is Nonrival

- Growth literature: Ideas are nonrival
  - Unlike rival goods, ideas are infinitely usable
- Data is another nonrival good
  - Clearly not a blueprint / recipe $\Rightarrow$ different from ideas
  - Ideas are production functions, data is a factor of production
  - Multiple engineers/algorithms can use same data at same time (within and across firms)
- Nonrivalry implies increasing returns to scale: $Y = F(D, X)$
  - Constant returns to rival inputs: $F(D, \lambda X) = \lambda F(D, X)$
  - Increasing returns to data and rival inputs: $F(\lambda D, \lambda X) > \lambda F(D, X)$

Data Property Rights Matter

- Key point: allocations with different degrees of data use

  $\Rightarrow$ different output, welfare, etc.

- How do different property rights affect the use of data?
  - "Firms own data" versus "consumers own data"

- To illustrate, we assume (plausibly?) the Coase theorem fails
  - Consumers can't commit to selling data to just one firm
  - Firms can't commit to not using data they acquire
  - Useful for showing the role of data sharing

Data is Nonrival $\Rightarrow$ Interesting Questions

- Do markets produce the right amount of data?
- Why don't firms (always) sell their data?
- Who should own data as it's created?
- Implications of data nonrivalry for antitrust, economic growth, and comparative advantage across countries?

We develop a framework for thinking through these questions

Outline

- Economic environment
- Allocations:
  - Optimal allocation
  - Firms own data
  - Consumers own data
  - Extreme privacy protection: outlaw data sharing
- Theory results and a numerical example

Basic Setup

Overview

- Representative consumer with a love for variety
- Innovation $\Rightarrow$ endogenous measure of varieties
- Nonrivalry of data $\Rightarrow$ increasing returns to scale
- How is data produced?
  - Learning by doing: each unit consumed $\rightarrow$ 1 unit of data
  - Alternative: separate PF (Tesla vs Google self-driving car)
- Any data equally useful in all firms $\Rightarrow$ one sector of economy
- Data depreciates fully each period

The Economic Environment

Utility	$\int_0^\infty e^{-\rho t} L_t \, u(c_t) dt$
Flow Utility	$u(c_t) = \log c_t$
Consumption per person	$c_t = \left(\int_0^{N_t} c_{it}^{\frac{\sigma-1}{\sigma}} di \right)^{\frac{\sigma}{\sigma-1}}$ with $\sigma > 1$
Data production	$J_{it} = c_{it} L_t$
Variety resource constraint	$c_{it} = Y_{it} / L_t$
Firm production	$Y_{it} = D_{it}^{\eta} L_{it}, \quad \eta \in (0,1)$
Data used by firm i	$D_{it} \leq \alpha x_{it} J_{it} + (1-\alpha) B_t$ (nonrivalry)
Data of firm i used by others	$D_{sit} \leq \tilde{x}_{it} J_{it}$
Data bundle	$B_t = \left(N_t^{-\frac{1}{\epsilon}} \int_0^{N_t} D_{sit}^{\frac{\epsilon-1}{\epsilon}} \, di \right)^{\frac{\epsilon}{\epsilon-1}}$ with $\epsilon > 1$
Innovation (new varieties)	$\dot{N}_t = \frac{1}{\chi} \cdot L_{et}$
Labor resource constraint	$L_{et} + \int_0^{N_t} L_{it} \, di = L_t$
Population growth (exogenous)	$L_t = L_0 e^{g_L t}$
Creative destruction	$\delta(\tilde{x}_{it}) = \frac{\delta_0}{2} \tilde{x}_{it}^2$ (equilibrium)

The Economic Environment: Simple Privacy Costs

Utility	$\int_0^\infty e^{-\rho t} L_t \, u(c_t, x_{it}, \tilde{x}_{it}) dt$
Flow Utility	$u(c_t, x_{it}, \tilde{x}_{it}) = \log c_t - \frac{\kappa}{2} \frac{1}{N_t^2} \int_0^{N_t} x_{it}^2 \, di - \frac{\tilde{\kappa}}{2} \frac{1}{N_t} \int_0^{N_t} \tilde{x}_{it}^2 \, di$
Consumption per person	$c_t = \left(\int_0^{N_t} c_{it}^{\frac{\sigma-1}{\sigma}} di \right)^{\frac{\sigma}{\sigma-1}}$ with $\sigma > 1$
Data production	$J_{it} = c_{it} L_t$
Variety resource constraint	$c_{it} = Y_{it} / L_t$
Firm production	$Y_{it} = D_{it}^{\eta} L_{it}, \quad \eta \in (0,1)$
Data used by firm i	$D_{it} \le \alpha x_{it} J_{it} + (1-\alpha) B_t$ (nonrivalry)
Data of firm i used by others	$D_{sit} \le \tilde{x}_{it} J_{it}$
Data bundle	$B_t = \left(N_t^{-\frac{1}{\epsilon}} \int_0^{N_t} D_{sit}^{\frac{\epsilon-1}{\epsilon}} \, di \right)^{\frac{\epsilon}{\epsilon-1}}$ with $\epsilon > 1$
Innovation (new varieties)	$\dot{N}_t = \frac{1}{\chi} \cdot L_{et}$
Labor resource constraint	$L_{et} + \int_0^{N_t} L_{it} \, di = L_t$
Population growth (exogenous)	$L_t = L_0 e^{g_L t}$
Creative destruction	$\delta(\tilde{x}_{it}) = \frac{\delta_0}{2} \tilde{x}_{it}^2$ (equilibrium)

The Planner Problem (using symmetry of firms)

$$\max_{\{L_{pt}, x_t, \tilde{x}_t\}} \int_0^\infty e^{-\tilde{\rho} t} L_0 \left(\log c_t - \frac{\kappa}{2} \frac{1}{N} x_t^2 - \frac{\tilde{\kappa}}{2} \tilde{x}_t^2 \right) dt, \quad \tilde{\rho} := \rho - g_L$$

subject to

$$
\begin{aligned}
c_t &= Y_t / L_t \\
Y_t &= N_t^{\frac{1}{\sigma-1}} D_{it}^{\eta} L_{pt} \\
D_{it} &= \alpha x_t Y_{it} + (1-\alpha) N_t \tilde{x}_t Y_{it} \\
Y_{it} &= D_{it}^{\eta} \cdot \frac{L_{pt}}{N_t} \\
\dot{N}_t &= \frac{1}{\chi} (L_t - L_{pt}) \\
L_t &= L_0 e^{g_L t}
\end{aligned}
$$

- More sharing ⇒ negative utility cost but more consumption
- Balance labor across production and entry/innovation

Scale Effect from Sharing Data

$$D_{it} = \alpha x_t J_{it} + (1-\alpha)\left(N_t^{-\frac{1}{\epsilon}} \int_0^{N_t} (\tilde{x}_t J_{it})^{\frac{\epsilon-1}{\epsilon}}\, di\right)^{\frac{\epsilon}{\epsilon-1}}$$

$$D_{it} = \alpha x_t Y_{it} + (1-\alpha) N_t \tilde{x}_t Y_{it}$$

$$= [\alpha x_t + (1-\alpha)\tilde{x}_t N_t] Y_{it}$$

- No sharing versus sharing:
  - No sharing: Only the αx_t term = no scale effect
  - Sharing: The $(1-\alpha)\tilde{x}_t N_t$ term = extra scale effect

  Source of Scale Effect: N_t scales with L_t

- Plugging into production function:

$$Y_{it} = \left([\alpha x_t + (1-\alpha)\tilde{x}_t N_t]^{\eta} L_{it}\right)^{\frac{1}{1-\eta}}$$

The Optimal Allocation on BGP (asymptotic)

$$\tilde{x}_{it} = \tilde{x}_{sp} = \left(\frac{1}{\tilde{\kappa}} \cdot \frac{\eta}{1-\eta}\right)^{1/2} \tag{1}$$

$$x_{it} = x_{sp} = \frac{\alpha}{1-\alpha}\frac{\tilde{\kappa}}{\kappa}\left(\frac{1}{\tilde{\kappa}} \cdot \frac{\eta}{1-\eta}\right)^{1/2} \tag{2}$$

$$L_{it}^{sp} = \chi\rho \cdot \frac{\sigma-1}{1-\eta} := \nu_{sp} \tag{3}$$

$$N_t^{sp} = \frac{L_t}{\chi\left(g_L + \nu_{sp}\right)} := \psi_{sp} L_t \tag{4}$$

$$L_{pt}^{sp} = \nu_{sp}\psi_{sp}L_t \tag{5}$$

$$Y_t^{sp} = \left(\nu_{sp}(1-\alpha)^{\eta}\tilde{x}_{sp}^{\eta}\right)^{\frac{1}{1-\eta}} \left(\psi_{sp}L_t\right)^{\frac{1}{\sigma-1}+\frac{1}{1-\eta}} \tag{6}$$

$$c_t^{sp} = \frac{Y_t}{L_t} = \left(\nu_{sp}(1-\alpha)^{\eta}\tilde{x}_{sp}^{\eta}\right)^{\frac{1}{1-\eta}} \left(\psi_{sp}L_t\right)^{\frac{1}{\sigma-1}+\frac{\eta}{1-\eta}} \tag{7}$$

$$g_c^{sp} = \left(\frac{1}{\sigma-1} + \frac{\eta}{1-\eta}\right) g_L \tag{8}$$

$$D_i^{sp} = \left((1-\alpha)\tilde{x}_{sp}\nu_{sp}\psi_{sp}L_t\right)^{\frac{1}{1-\eta}} \tag{9}$$

$$D^{sp} = ND_i = \left((1-\alpha)\tilde{x}_{sp}\nu_{sp}\right)^{\frac{1}{1-\eta}} \left(\psi_{sp}L_t\right)^{1+\frac{1}{1-\eta}} \tag{10}$$

$$Y_{it}^{sp} = \left(\nu_{sp}(1-\alpha)^{\eta}\tilde{x}_{sp}^{\eta}\right)^{\frac{1}{1-\eta}} \left(\psi_{sp}L_t\right)^{\frac{\eta}{1-\eta}} \tag{11}$$

The Optimal Allocation: GDP per person

$$c_t^{sp} = \frac{Y_t}{L_t} = \left(\nu_{sp}(1-\alpha)^{\eta}\tilde{x}_{sp}^{\eta}\right)^{\frac{1}{1-\eta}} \left(\psi_{sp}L_t\right)^{\frac{1}{\sigma-1}+\frac{\eta}{1-\eta}}$$

$$g_c^{sp} = \left(\frac{1}{\sigma-1} + \frac{\eta}{1-\eta}\right) g_L$$

- Scale effect: $\underbrace{\frac{1}{\sigma-1}}_{\text{Love of Variety}} + \underbrace{\frac{\eta}{1-\eta}}_{\text{Data}}$
- More people make more data and all firms use all shared data

The Optimal Allocation: Data, Firm Size, Variety

$$\tilde{x}_{sp} = \left(\frac{1}{\tilde{\kappa}} \cdot \frac{\eta}{1-\eta} \right)^{1/2}$$

$$L_{it}^{sp} = \chi\rho \cdot \frac{\sigma - 1}{1-\eta} := \nu_{sp}$$

$$N_t^{sp} = \frac{L_t}{\chi g_L + \nu_{sp}} := \psi_{sp} L_t$$

- Data shared increasing in data production elasticity and decreasing in privacy cost
- Firm size constant on BGP. N has opposite comparative statics
- Higher entry cost, time preference, population growth, and elasticity of substitution raise firm size and reduce varieties
- Higher η raises firm size and reduces varieties: Entry does not create data

Firms Own Data

Firms Own Data: Consumer Problem

- Firms own data and choose one data policy $(x_{it}, \tilde{x}_{it})$ applied to all consumers

- Consumers just choose consumption:

$$
\begin{aligned}
U_0 &= \max_{\{c_{it}\}} \int_0^{\infty} e^{-\tilde{\rho} t} L_0 u(c_t, x_{it}, \tilde{x}_{it}) dt \\
\text{s.t.} \quad c_t &= \left(\int_0^{N_t} c_{it}^{\frac{\sigma-1}{\sigma}} di \right)^{\frac{\sigma}{\sigma-1}} \\
\dot{a}_t &= (r_t - g_L) a_t + w_t - \int_0^{N_t} p_{it} c_{it} di
\end{aligned}
$$

Firms own Data: Data Decisions

- Firms buy D_{bit} data from intermediary at given price p_b
- Firms sell D_{sit} data to intermediary at chosen price p_{si}
  - Perfect competition inconsistent with nonrival data!
  - Monopolistically competitive with own data
  - See the intermediary's downward-sloping demand curve and set price
- How much data to use / sell?
  - x_{it}: Use all of own data $\Rightarrow x_{it} = 1$
  - $\tilde{x}_{it}$: Trade off = selling data versus creative destruction

    $\delta(\tilde{x}_{it})$ = Poisson rate transferring ownership of variety

Firms own the Data: Incumbent Firm Problem

- Monopolistically competitive firm takes demand for variety as given (from FOC of consumer problem): $p_{it} = \left(\frac{c_t}{c_{it}}\right)^{\frac{1}{\sigma}} = \left(\frac{Y_t}{Y_{it}}\right)^{\frac{1}{\sigma}}$

$$r_t V_{it} = \max_{L_{it}, D_{bit}, x_{it}, \tilde{x}_{it}} \left(\frac{Y_t}{Y_{it}}\right)^{\frac{1}{\sigma}} Y_{it} - w_t L_{it} - p_{bt} D_{bit} + p_{st} \tilde{x}_{it} Y_{it} + \dot{V}_{it} - \delta(\tilde{x}_{it}) V_{it}$$

$$\text{s.t.} \quad Y_{it} = D_{it}^{\eta} L_{it}$$

$$D_{it} = \alpha x_{it} Y_{it} + (1-\alpha) D_{bit}$$

$$x_{it} \in [0,1], \; \tilde{x}_{it} \in [0,1]$$

$$p_{sit} = \lambda_{DI} N_t^{-\frac{1}{\epsilon}} \left(\frac{B_t}{\tilde{x}_{it} Y_{it}}\right)^{\frac{1}{\epsilon}}$$

- Data Intermediary $(p_{bt}, p_{st}, D_{bit})$ and Free Entry complete eqm.

Firms own the Data: Data Intermediary Problem

- A monopolist takes data purchase price as given and sees the downward sloping demand curve for data $p_{bt}(D_{bit})$:

$$\max_{p_{bt}, D_{sit}} \quad p_{bt} \int_0^{N_t} D_{bit}\, di - p_{st} \int_0^{N_t} D_{sit}\, di$$

s.t.

$$D_{bit} \leq B_t = \left(N_t^{-\frac{1}{\epsilon}} \int_0^{N_t} D_{sit}^{\frac{\epsilon-1}{\epsilon}}\, di \right)^{\frac{\epsilon}{\epsilon-1}}$$

$$p_{bt} \leq p_{bt}^*$$

- Free entry at zero cost $\Rightarrow$ zero profits
- Problem incorporates data nonrivalry
  - Buys data once from each firm
  - But can sell the same bundle multiple times

Entry: Innovation Creates a New Variety

- χ units of labor needed to create an additional variety
- Free entry condition:

$$\chi w_t = V_{it} + \frac{\int_0^{N_t} \delta(\tilde{x}_{it}) V_{it} \, di}{\dot{N}_t}$$

- The value of a new variety and the per-entrant share of business stealing from creative destruction

Firms Own Data: A “No Trade” Law

- What if the government, in an attempt to protect consumers privacy, makes data sharing illegal?
- Government chooses
  - $x_{it} \in (0, 1]$
  - $\tilde{x}_{it} = 0$
- We call this the “Outlaw Sharing” allocation

Consumers Own Data

Consumers own Data: Consumer Problem

- Consumers own data, so now choose how much to sell $(x_{it}, \tilde{x}_{it})$:

$$U_0 = \max_{\{c_{it}, x_{it}, \tilde{x}_{it}\}} \int_0^\infty e^{-\tilde{\rho}t} L_0 u(c_t, x_{it}, \tilde{x}_{it}) dt$$

s.t. $$c_t = \left(\int_0^{N_t} c_{it}^{\frac{\sigma-1}{\sigma}} di \right)^{\frac{\sigma}{\sigma-1}}$$

$$\dot{a}_t = (r_t - g_L)a_t + w_t - \int_0^{N_t} p_{it} c_{it} di + \int_0^{N_t} x_{it} p_{st}^a c_{it} di + \int_0^{N_t} \tilde{x}_{it} p_{st}^b c_{it} di$$

- Firm problem similar to before, but now takes $x, \tilde{x}$ as given, can't sell data, and has to buy "own" data

Consumers own the Data: Incumbent Firm Problem

- Monopolistically competitive firm takes demand for variety as given (from FOC of consumer problem):
$q_{it} = \left(\frac{c_t}{c_{it}}\right)^{\frac{1}{\sigma}} = \left(\frac{Y_t}{Y_{it}}\right)^{\frac{1}{\sigma}} = p_{it} - x_{it}p^a_{st} - \tilde{x}_{it}p^b_{st}$
- Firm buys data on its own variety (D_{ait}) and data on other firms varieties (D_{bit})

$$
\begin{aligned}
r_t V_{it} = \max_{L_{it}, D_{ait}, D_{bit}} & \left[\left(\frac{Y_t}{Y_{it}}\right)^{\frac{1}{\sigma}} + p^a_{st} x_{it} + p^b_{st} \tilde{x}_{it} \right] Y_{it} - w_t L_{it} \\
& - p_{at} D_{ait} - p_{bt} D_{bit} + \dot{V}_{it} - \delta(\tilde{x}_t) V_{it}
\end{aligned}
$$

$$
\begin{aligned}
\text{s.t.} \quad Y_{it} &= D^{\eta}_{it} L_{it} \\
D_{it} &= \alpha D_{ait} + (1-\alpha) D_{bit} \\
D_{ait} &\geq 0, \quad D_{bit} \geq 0
\end{aligned}
$$

Key Forces: Consumers vs. Firms vs. Outlaw Sharing

- Firms
  - use all data on own variety, ignoring consumer privacy
  - restrict data sharing because of creative destruction
- Consumers
  - respect their own privacy concerns
  - sell data broadly, ignoring creative destruction
- Outlaw sharing
  - maximizes privacy gains
  - missing scale effect reduces consumption

Results: Comparing Allocations

1. Planner Problem
2. Firms Own Data
3. Outlaw Data Sharing
4. Consumers Own Data

Key Allocations: $alloc \in \{sp, f, c, ns\}$

- Firm size: $L_i^{alloc} = L_{pt}/N_t = \nu_{alloc}$

$$\nu_{sp} := \chi\rho \cdot \frac{\sigma - 1}{1 - \eta}$$
$$\nu_{os} := \chi\rho \cdot \frac{\sigma - 1}{1 - \sigma\eta}$$
$$\nu_{c} := \chi g_L \cdot \frac{\rho + \delta(\tilde{x}_c)}{g_L + \delta(\tilde{x}_c)} \cdot \frac{\sigma - 1}{1 - \sigma\eta}$$
$$\nu_{f} := \chi g_L \cdot \frac{\rho + \delta(\tilde{x}_f)}{g_L + \delta(\tilde{x}_f)} \cdot \frac{\sigma - 1}{1 - \sigma\eta\frac{\epsilon - 1}{\epsilon}}$$

- Number of firms: $N_t^{alloc} = \psi_{alloc} L_t$

$$\psi_{alloc} := \frac{1}{\chi g_L + \nu_{alloc}}$$

Output

- For $alloc \in \{sp, c, f\}$:

$$Y_t^{alloc} = \left[\nu_{alloc}(1-\alpha)^{\eta}\tilde{x}_{alloc}^{\eta}\right]^{\frac{1}{1-\eta}} (\psi_{alloc}L_t)^{1+\frac{1}{\sigma-1}+\frac{\eta}{1-\eta}}$$

- For Outlaw Sharing:

$$Y_t^{os} = \left[\nu_{os}\alpha^{\eta}x_{os}^{\alpha}\right]^{\frac{1}{1-\eta}} (\psi_{os}L_t)^{1+\frac{1}{\sigma-1}}$$

- Two source of increasing returns to scale:
  - Standard variety effect: $\frac{\sigma}{\sigma-1}$
  - Data sharing: $\frac{\eta}{1-\eta}$
- Recall $\tilde{x}_t > 0$ from data sharing $\Rightarrow$ scale effect

Data Sharing

Own Firm Data	Sharing with Other Firms
$x_{sp} = \frac{\alpha}{1-\alpha} \frac{\tilde{\kappa}}{\kappa} \left(\frac{1}{\tilde{\kappa}} \cdot \frac{\eta}{1-\eta} \right)^{1/2}$	$\tilde{x}_{sp} = \left(\frac{1}{\tilde{\kappa}} \cdot \frac{\eta}{1-\eta} \right)^{1/2}$
$x_f = 1$	$\tilde{x}_f = \left(\frac{2\Gamma\rho}{(2-\Gamma)\delta_0} \right)^{1/2}, \Gamma := \frac{\eta(\sigma-1)}{\frac{\epsilon}{\epsilon-1} - \sigma\eta}$
$x_{os} \in (0, 1]$	$\tilde{x}_{os} = 0$
$x_c = \frac{\alpha}{1-\alpha} \frac{\tilde{\kappa}}{\kappa} \left(\frac{1}{\tilde{\kappa}} \cdot \frac{\eta}{1-\eta} \cdot \frac{\sigma-1}{\sigma} \right)^{1/2}$	$\tilde{x}_c = \left(\frac{1}{\tilde{\kappa}} \cdot \frac{\eta}{1-\eta} \cdot \frac{\sigma-1}{\sigma} \right)^{1/2}$

- Firms fear creative destruction and share less than planner (δ_0)
- Consumers share less than planner because of mark up
- No sharing law restricts data even more
- Firms use more own-variety data compared to consumer/planner

Numerical Example: Parameter Values

Description	Parameter	Value
Importance of data	η	0.06
Elasticity of substitution	σ	4
Weight on privacy	$\kappa = \tilde{\kappa}$	0.20
Population level	L_0	100
Population growth rate	g_L	0.02
Rate of time preference	ρ	0.025
Labor cost of entry	χ	0.01
Creative destruction	δ_0	0.4
Weight on own data	α	1/2
Elasticity of Substitution (data)	ϵ	50
Use of own data in NS	$\bar{x}$	1

Numerical Example: How large is η? (Approach 1 - Data Share)

- Share of GDP spent on data = $\frac{\eta}{1-\eta}\frac{\sigma-1}{\sigma}$
- Similar formula/quantity when consumers or firms own data
- Set $\sigma = 4$
- If data share of GDP is 5% $\Rightarrow \eta = 0.0625$
- If data share of GDP is 10% $\Rightarrow \eta = 0.12$
- Approach will be to explore $\eta \in \{0.03, 0.06, 0.12\}$

Numerical Example: Consumption Equivalent Welfare

$$U_{ss}^{alloc} = \frac{1}{\tilde{\rho}} \left(\log c_0^{alloc} - \frac{\tilde{\kappa}}{2} \tilde{x}_{alloc}^2 + \frac{g_c^{alloc}}{\tilde{\rho}} \right).$$

Let $U_{ss}^{alloc}(\lambda)$ denote steady-state welfare when we perturb the allocation of consumption by some proportion λ:

$$U_{ss}^{alloc}(\lambda) = \frac{1}{\tilde{\rho}} \left(\log(\lambda c_0^{alloc}) - \frac{\tilde{\kappa}}{2} \tilde{x}_{alloc}^2 + \frac{g_c^{alloc}}{\tilde{\rho}} \right).$$

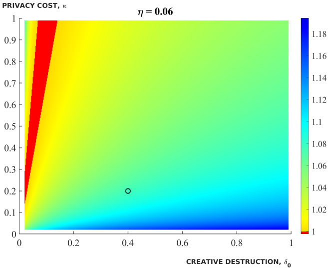
Define consumption equivalent welfare as λ^{alloc}:

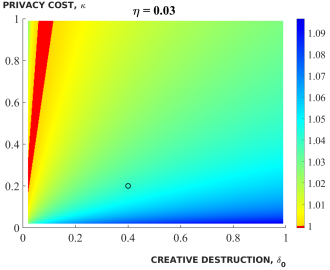
$U_{ss}^{sp}(\lambda^{alloc}) = U_{ss}^{alloc}(1)$ with

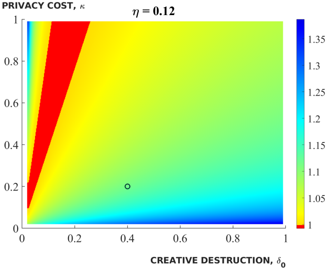
$$\log \lambda^{alloc} = \underbrace{\log c_0^{alloc} - \log c_0^{sp}}_{\text{Level term}} - \underbrace{\frac{\tilde{\kappa}}{2} \left(\tilde{x}_{alloc}^2 - \tilde{x}_{sp}^2 \right)}_{\text{Privacy term}} + \underbrace{\frac{g_c^{alloc} - g_c^{sp}}{\tilde{\rho}}}_{\text{Growth term}}$$

Note: The x_{it} terms drop out because scaled by $1/N$

Welfare Sensitivity Analysis (η, δ, κ): λ^c/λ^f







Allocations: Baseline

Allocation	Data Sharing “own” x	“others” $\tilde{x}$	Firm size ν	Variety $N/L = \psi$	Consu-mption c	Growth g	Creative Destruct. δ
Social Planner	0.66	0.66	1304	665	18.6	0.67%	0.0870
Consumers Own Data	0.59	0.59	1482	594	18.3	0.67%	0.0696
Firms Own Data	1	0.16	1838	491	16.0	0.67%	0.0052
Outlaw Sharing	1	0	2000	455	7.3	0.50%	0

- Firms overuse their own data and undershare with others
- Consumers share less data than planner, but not by much
- Growth rate scale effect is modest, level differences are large

Consumption Equivalent Welfare

Allocation	Welfare λ	$\log\lambda$	Level term	Privacy term	Growth term
Optimal Allocation	1	0	..	..	..
Consumers Own Data	0.9886	-0.0115	-0.0202	0.0087	0.0000
Firms Own Data	0.8917	-0.1146	-0.1555	0.0409	0.0000
Outlaw Sharing	0.3429	-1.0703	-0.9399	0.0435	-0.1739

- Outlaw sharing: particularly harmful law (66 percent worse!)
- Firms own data: substantially lower welfare (11 percent worse)
- Consumers own data: nearly optimal (1 or 2 percent worse)

Implications for IO

- Firms that use data might grow fast compared to those that don't
- Firms would like to merge into one single economy-wide firm
  - Implications for antitrust
  - Price/quantity behavior
- What are the costs of forced sharing?
  - Disincentive to collect/create data
  - Data as a barrier to entry
    (extension to quality ladder model)
  - Markets unraveling
- Targeted mandatory sharing?
  - E.g., airplane safety (after a crash)

Data versus Ideas: Excludability

- Maybe technologically easier to transmit data than ideas (usb key vs. education) . . .
- But data can be encrypted and monitored
- Data seems highly excludable
  - Idea: use machine learning to train self-driving car algorithm
  - ML needs lots of data. Each firm gathering own data

The Boundaries of Data Diffusion: Firms and Countries

- How does data diffuse across firms and countries?
  - Ideas eventually diffuse across firms or countries, so no country scale effect (e.g. HK vs China)
  - What about data?
- Scale effects and country size
  - Larger countries may have an important advantage as data grows in importance
- Scale effects and institutions
  - What if China mandates data sharing across Chinese firms and U.S. has no such policy
  - What if consumers in China have different privacy concerns than in the U.S. or Europe?

Conclusion

- Nonrival data ⇒ large social gain from sharing data
- If firms own data, they may:
  - privately use more data than consumers/planner would
  - sell less data across firms than consumers/planner would
- Nonrivalry ⇒ Laws that outlaw sharing could be very harmful
- Consumers owning data good at balancing privacy and sharing